



AI 9 OIPE

RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/804,481

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

ENTERED

4 <110> APPLICANT: de Graaf, David
 5 Lander, Eric S.
 7 <120> TITLE OF INVENTION: Novel Small Nuclear RNA Vectors and Uses
 8 Therefor
 10 <130> FILE REFERENCE: 2825.1023-001
 12 <140> CURRENT APPLICATION NUMBER: 09/804,481
 13 <141> CURRENT FILING DATE: 2001-03-12
 15 <150> PRIOR APPLICATION NUMBER: 60/188,304
 16 <151> PRIOR FILING DATE: 2000-03-10
 18 <160> NUMBER OF SEQ ID NOS: 11
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 4639
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 27 <400> SEQUENCE: 1
 28 gacggatcgg gagatctccc gatccctat ggtcgactct cagtacaatc tgctctgatg 60
 29 ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
 30 cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgagct cggtagcccg 180
 31 ggagatccgg taaggaccag cttctttggg agagaacaga cgcaggggcg ggagggaaaa 240
 32 agggagaggc agacgtcaact tccccttggc ggctctggca gcagattggt cggttgagtg 300
 33 gcagaaaggc agacggggac tgggcaaggc actgtcggtg acatcacgga cagggcgact 360
 34 tctatgtaga tgaggcagcg cagaggctgc tgcttcgcca cttgctgctt caccacgaag 420
 35 gagttcccggt gccctgggag cgggttcagg accgctgac ggaagtgaga atcccagctg 480
 36 tgtgtcaggg ctggaaggcg ctcgggagtg cgcggggcaa gtgaccgtgt gtgtaaagag 540
 37 tgaggcgtat gaggtgtgt cggggcagag gcccaagatc tcaagggccc ataacatgtg 600
 38 taccatcgat tgcaggggag ataocatgat cacgaagggtg gttttccag ggcgaggctt 660
 39 atccattgca ctccggatgt gctgacccct gcgatttccc caaagcttg aaactcgact 720
 40 gcataatttg tggtagtggg ggaactgctt cgcgctttcc cctgactttc tggagtttca 780
 41 aaagtagact gtacgctaac cggatccctc agagtcgacc tgcaggcatg cagaagacaa 840
 42 ttagcaggca tgctgggat gcggtgggct ctatggcttc tgaggcggaa agaaccagct 900
 43 ggggctctag ggggtatccc cagcgccct gtagcggcgc attaagcgc gcgggtgtgg 960
 44 tggttacgcg cagcgtgacc gctacacttg ccagcgcct agcgcgcct cctttcgctt 1020
 45 tcttcccttc ctttctcgcc acgttcgccc gctttccccc tcaagctcta aatcggggca 1080
 46 tccctttagg gttccgattt agtgctttac ggcaacctga ccccaaaaaa cttgattagg 1140
 47 gtgatggttc acgtagtggg ccacgcacct gatagacggt ttttcgacct ttgacgttg 1200
 48 agtccacgtt ctttaatatg ggaactctgt tccaaactgg aacaacactc aaccctatct 1260
 49 cggctctatc ttttgattta taagggattt tggggatttc ggcctatttg ttaaaaaatg 1320
 50 agctgattta acaaaaattt aacgcgaatt aattctgtgg aatgtgtgtc agttagggtg 1380
 51 tggaaagtcc ccaggctccc caggcaggca gaagtatgca aagcatgcat ctcaattagt 1440
 52 cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg caaagcatgc 1500
 53 atctcaatta gtcagcaacc atagtcccg ccctaactcc gcccatcccg cccctaactc 1560
 54 cgcccagttc cgccattct ccgcccagtg gctgactaat tttttttatt tatgcagagg 1620

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/804,481

DATE: 01/27/2002

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

55 ccgaggccgc ctctgcctct gagctattcc agaagtagtg aggaggcttt tttggaggcc 1680
56 taggcttttg caaaaagctc ccgggagctt gtatatccat tttcggatct gatcagcacg 1740
57 tggtgacaat taatcatcgg catagtatat cggcatagta taatacgaca aggtgaggaa 1800
58 ctaaaccatg gccaaagtga ccagtgcgtt tccggtgctc accgcgcgcg acgtcgcgcg 1860
59 agcggtcgag ttctggaccg accggctcgg gttctcccg gacttcgttg aggacgactt 1920
60 cgcgggtgtg gtcggggaac acgtgaccct gttcatcagc gcgggtccagg accagggtgt 1980
61 gccggacaac accctggcct ggggtgtggg gcgcggcctg gacgagctgt acgccgagt 2040
62 gtcggagggtc gtgtccacga acttccggga cgcctccggg ccggccatga ccgagatcgg 2100
63 cgagcagccg tgggggcggg agttcgcctt gcgcgaccgg gccggcaact gcgtgcactt 2160
64 cgtggccgag gagcaggact gacacgtgct acgagatttc gattccaccg ccgccttcta 2220
65 tgaaagggtt ggcttcggaa tcgttttccg ggacgcgggc tggatgatcc tccagcgcgg 2280
66 ggatctcatg ctggagttct tcgcccaccc caacttggtt attgcagctt ataattggtta 2340
67 caaataaagc aatagcatca caaatttcac aaataaagca tttttttcac tgcattctag 2400
68 ttgtggtttg tccaaactca tcaatgtatc ttatcatgtc tgtataccgt cgacctctag 2460
69 ctagagcttg gcgtaatcat ggtcatagct gtttctctgt tgaaattggt atccgctcac 2520
70 aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctgggggt cctaattgag 2580
71 gagctaactc acattaattg cgttgcgctc actgcccgtt ttccagtcgg gaaacctgtc 2640
72 gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc gtattggcg 2700
73 ctcttccgct tcctcgtcga ctgactcgtt gcgctcggtc gttcggctgc ggcgagcgg 2760
74 atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa 2820
75 gaacatgtga gcaaaaggcc agcaaaaggc caggaaacct aaaaaggcgg cgttgctggc 2880
76 gtttttccat aggtccgcgc cccctgacga gcatcacaaa aatcgacgct caagtcagag 2940
77 gtggcgaaac ccgacaggac tataaagata ccaggcggtt ccccttgga gctccctcgt 3000
78 gcgctctcct gttccgaccc tgccgcttac cggataacct tccgccttcc tcccttcggg 3060
79 aagcgtggcg ctttctcaat gctcacgctg taggtatctc agttcgggtg aggtcgttcg 3120
80 ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg 3180
81 taactatcgt cttgagttca acccggttag acacgaacta tcgccactgg cagcagccac 3240
82 tggtaacagg attagcagag cgaggtaggt aggcggtgct acagagttct tgaagtgggt 3300
83 gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt 3360
84 taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg 3420
85 tgggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc 3480
86 tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggtattt 3540
87 ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt 3600
88 taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag 3660
89 tgaggcaact atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt 3720
90 cgtgtagata actacgatac gggagggtt accatctggc cccagtgtc caatgatacc 3780
91 gcgagaccca cgtcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc 3840
92 cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg 3900
93 ggaagctaga gtaagtgtt cggcagttaa tagtttgccg aacgttggtt ccattgctac 3960
94 aggcacgtg gtgtcacgct cgtogtttg tatggcttca ttcagctccg gttcccaacg 4020
95 atcaaggcga gttacatgat ccccattgtt gtgcaaaaa gcggttagct ccttcggtcc 4080
96 tccgactcgt tcgagaagta agttggccgc agtggtatca ctcagtgtta tggcagcact 4140
97 gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc 4200
98 aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat 4260
99 acgggataat accgcgccac atagcagaac tttaaaagt ctcattcatt gaaaacgttc 4320
100 ttccggggcg aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaaccac 4380
101 tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa 4440
102 aacaggaagg caaatgccg caaaaaagg aataaggcg acacggaaat gttgaatact 4500
103 catactcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg 4560

```

RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/804,481

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

104 atacatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca catttccccg 4620
105 aaaagtgccca cctgacgtc 4639
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 5
109 <212> TYPE: DNA
110 <213> ORGANISM: Artificial Sequence
112 <220> FEATURE:
113 <223> OTHER INFORMATION: single-stranded restriction fragment overhand
115 <400> SEQUENCE: 2
116 gcagg 5
118 <210> SEQ ID NO: 3
119 <211> LENGTH: 5
120 <212> TYPE: DNA
121 <213> ORGANISM: Artificial Sequence
123 <220> FEATURE:
124 <223> OTHER INFORMATION: single-stranded restriction fragment overhang
126 <400> SEQUENCE: 3
127 tgaga 5
129 <210> SEQ ID NO: 4
130 <211> LENGTH: 33
131 <212> TYPE: DNA
132 <213> ORGANISM: Artificial Sequence
134 <220> FEATURE:
135 <223> OTHER INFORMATION: recognition site
137 <221> NAME/KEY: misc_feature
138 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 22, 23, 24,
139 25, 26, 27, 28, 29, 30, 31, 32, 33
140 <223> OTHER INFORMATION: n = A,T,C or G
142 <400> SEQUENCE: 4
143 nnnnnnnnnn acnnngtay cnnnnnnnnn nnn 33
145 <210> SEQ ID NO: 5
146 <211> LENGTH: 33
147 <212> TYPE: DNA
148 <213> ORGANISM: Artificial Sequence
150 <220> FEATURE:
151 <223> OTHER INFORMATION: recognition site
153 <221> NAME/KEY: misc_feature
154 <222> LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 19,
155 20, 21, 27, 28, 29, 30, 31, 32, 33
156 <223> OTHER INFORMATION: n = A,T,C or G
158 <400> SEQUENCE: 5
159 nnnnnnnnnn nnnntgnnn ncatrgnnnn nnn 33
161 <210> SEQ ID NO: 6
162 <211> LENGTH: 10
163 <212> TYPE: DNA
164 <213> ORGANISM: Artificial Sequence
166 <220> FEATURE:
167 <223> OTHER INFORMATION: modification fragment
169 <400> SEQUENCE: 6

```

RAW SEQUENCE LISTING

DATE: 01/27/2002

PATENT APPLICATION: US/09/804,481

TIME: 15:27:49

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

```

170 cacaaacaca 10
172 <210> SEQ ID NO: 7
173 <211> LENGTH: 12
174 <212> TYPE: DNA
175 <213> ORGANISM: Artificial Sequence
177 <220> FEATURE:
178 <223> OTHER INFORMATION: modification fragment
180 <400> SEQUENCE: 7
181 tccacaaaca ca 12
183 <210> SEQ ID NO: 8
184 <211> LENGTH: 15
185 <212> TYPE: DNA
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: modification fragment
191 <400> SEQUENCE: 8
192 tcgtccacaa acaca 15
194 <210> SEQ ID NO: 9
195 <211> LENGTH: 12
196 <212> TYPE: DNA
197 <213> ORGANISM: Artificial Sequence
199 <220> FEATURE:
200 <223> OTHER INFORMATION: modification fragment
202 <400> SEQUENCE: 9
203 cacaaacaca ac 12
205 <210> SEQ ID NO: 10
206 <211> LENGTH: 10
207 <212> TYPE: DNA
208 <213> ORGANISM: Artificial Sequence
210 <220> FEATURE:
211 <223> OTHER INFORMATION: modification fragment
213 <400> SEQUENCE: 10
214 cacaaacacg 10
216 <210> SEQ ID NO: 11
217 <211> LENGTH: 59
218 <212> TYPE: DNA
219 <213> ORGANISM: Artificial Sequence
221 <220> FEATURE:
222 <223> OTHER INFORMATION: vector construct
224 <400> SEQUENCE: 11
225 ggcccaagat ctcaagggcc cataacatgt gtaccatcga ttgcagggga gataccatg 59

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/804,481

DATE: 01/27/2002

TIME: 15:27:50

Input Set : A:\28251023001.TXT

Output Set: N:\CRF3\01272002\I804481.raw

L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5